Deterrence in the Cold War and the "War on Terror"[1]

David K. Levine

Robert A. Levine[2]

First Version: January 23, 2006 This Version: September 1, 2005

[2] Department of Economics, UCLA, and Canyon Analysts.

The only thing we have to fear is fear itself.

-- Franklin Delano Roosevelt, First Inaugural Address, 1933

1. Background

The authors of this paper, father and son, span Jack Hirshleifer's career, and his interests, in economic theory and in its application to issues of public policy. Robert Levine was an undergraduate student in Jack's section of Alvin Hansen's macroeconomics course at Harvard in the 1940s. He went on to a career analyzing economic and strategic policy in and out of government; he was Jack's colleague at the RAND Corporation in the 1950s and 1960s. David Levine was Jack's student as a UCLA undergraduate in the 1970s and became his colleague on the UCLA economics faculty in the 1980s, specializing in economic theory, particularly game theory.

We have chosen to write this paper about the role of economics in understanding and making military policy, a subject of great interest to Jack throughout his career, which lasted from the start of the cold war to the "war on terror". We focus on the issue of deterrence – critical to the cold war – and ask what changes in the theory and applications are needed to cope with terror.

Our central conclusions for future policy are that: deterrence of terrorist states will resemble cold-war deterrence; deterring terrorist groups, however, will be more difficult; but that in either case, failure of deterrence should have far less traumatic consequences than during the cold war, unless we are overcome by fear itself.

Economic analysis entered cold war thinking and policymaking in two major ways, one conventional, the other more innovative. The first was the examination of the costs and benefits of alternative force postures, weapons, and other choices. The second was the application of game theory to strategy through deterrence.[3] Here our interest is in the latter.

The entry of game theory into U.S. strategy via sophisticated concepts of deterrence was innovative and intricate. For the first years of the cold war, the concept of deterrence was simple: the threat of "massive retaliation" with nuclear weapons a la Hiroshima and Nagasaki would cause the "opponent", the Soviet Union, to refrain from

---

[3] Jack Hirshleifer contributed to both. As an employee and consultant at RAND, he had 45 internal and external publications to his name. The most relevant are Hirshleifer [1949, 1963, 1989].

aggression. Even after the Soviets built their own nuclear and thermonuclear weapons, Americans believed that our superiority would suffice to deter them. But American superiority fell into doubt with Soviet achievement of the first space launch in 1957, and with subsequent fears of the "missile gap". Even before that, however the indiscriminate nature of the city-busting that had helped end World War II was criticized on grounds not only of immorality but of military effectiveness—the Russians did not base their retaliatory weapons in Moscow and Leningrad.

Beginning in the 1950s, economist Thomas Schelling applied game theory to deterrence to generate a far more discriminating theory of deterrence as bargaining between adversaries (and partners, who could not always be distinguished from adversaries) involving their common interest in avoiding nuclear war as well as their opposing interests.[4] The outputs were not only unilateral military strategies but also written and implicit arms control agreements. Schelling won the 2005 Nobel Prize in economics, largely for his work of the 1950s and 1960s.

Schelling was the central theorist; others like mathematician Herman Kahn and mathematical logician Albert Wohlstetter applied the theory to the world of the 1950s and 1960s. Kahn's massive and eclectic book, *On Thermonuclear War,* distinguished two types of deterrence—deterring an enemy first nuclear strike, and using the threat of our own first strike to deter lesser aggression—and quantified the horrifying results of various possible wars. He proposed policies like civil defense that he argued would make nuclear wars less likely and less destructive[5]. Critics accused him of thinking about the unthinkable, so he made that the title of his next book.[6] Wohlstetter applied deterrence theory to the structure of U.S. nuclear forces, initially as a consultant to the Department of Defense, and then in a long series of writings relating specific weapons decisions and military postures to avoiding nuclear war, and to achieving political objectives such as the preservation of west European freedom.[7]

Applied theory was still theory, but President Kennedy's 1961 appointment of Robert McNamara as Secretary of Defense brought the theory to the center of American strategic policy. Schelling, Kahn, and Wohlstetter never joined the government except as

[4] This was encompassed initially in two books, Schelling [1960] and Schelling and Halperin [1961.
[5] Kahn [1961a].
[6] Kahn [1962].
[7] Kahn [1961b].

consultants, but many of their colleagues and disciples became members of McNamara's band of "whiz kids". The Secretary quickly recognized the importance of deterrent detail: massive retaliation was replaced as the center of U.S. strategy by "counterforce" targeting to weaken the enemy's retaliatory capability, and "flexible response" to different Soviet threats at different military levels, always with the possibility of escalation to higher levels. Controversy, some of it near-theological, abounded, particularly over whether making deterrence more realistic increased the possibility of nuclear war stemming from the implementation of deterrent threats; most of the controversy was in the Schelling context. And the new doctrines remained in place throughout the cold war.

Deterrence worked—or at least it did not not work; neither nuclear war nor major Soviet aggression took place—but it worked under very special circumstances. Schelling stressed that the theory "is based on the assumption that the participants coolly and 'rationally' calculate their advantages according to a consistent value system."[8] He then went on to list varieties of "irrationality," [discussed further below], but Schelling rationality applied to the Soviet Union as well as the U.S. Although a few Sovietologists pointed out that the Russians did not think about deterrence in the way we did, the theory did not require them to. The USSR was a large state with many Schellingesque "advantages", and it was conservative in the sense that its leaders wanted to conserve those advantages as well as their own power. They understood what could be lost in a nuclear war as well as what could be gained by risking one.

The "War on Terror" has changed the demands on deterrence, and its applicability, more radically in several ways. The notion that terrorists "coolly and 'rationally' calculate their advantages according to a consistent value system is an implausible one; and the stakes of the game have dropped sharply. Additionally, cold war nuclear deterrence was a two-player game. The French nuclear *force de frappe,* which President Charles de Gaulle insisted was necessary to maintain strategic independence, and the equally small Chinese nuclear force, complicated matters but not very much.

By way of contrast, extremist Muslim terrorism is diffuse, with no central decision-making locus like Moscow in the cold war. Rhetoric frequently characterizes Islamic terrorism as "al-Qaeda" or even personalizes it down to Osama bin Laden and/or

---

[8] Schelling [1960] p.16.

his supposed associate in Iraq, Abu Musad al-Zarkawi, but in fact al-Qaeda is at most a loose network, not a command structure in which a central decision influenced by deterrence can be transmitted down to the branches. Capturing or killing bin Laden or al-Zarkawi would be a setback but far from a death knell for the organization. And while al-Qaeda is probably the largest of terrorist networks, it is far from the only one. Deterrence of terrorists must therefore cope with a multitude of independent or autonomous centers.

Further, the centers vary widely in the benefits they ascribe to their activities and the costs they are willing to bear—indeed even what they consider costs, as witness the ubiquity of suicide bombers. Some so-called terrorists are essentially thugs—the late Yugoslav commander with the *nom de guerre* Arkan was clearly that, and the entire Yugoslav government at that time has been characterized that way; so have some elements of the Irish Republican Army. Some terrorists in Baghdad, whose ideology forms a thin cover for extortion, looting, and theft, also fall into that category. Ironically, such groups seem the most likely to make cost-benefit calculations that deterrence can affect.

At the other pole, however, are the large number of terrorist groups that burn with pure ideology—hatred for the infidels. The 9/11 hijackers seem to fit into that category; how could they have been deterred? Between the poles are groups with more specific objectives—getting the Americans out of Iraq or the Sunnis back into business, getting the Israelis out of the West Bank or out of Israel. Their benefits are clearer, their costs at least partially calculable. Although deterrence may be applied to some of these cases, no single model similar to that of the cold war can cover them all.

The second critical difference between the cold war and war on terrorism is that the threat level has decreased by orders of magnitude. In a notorious table entitled "Tragic but Distinguishable Postwar States," Herman Kahn listed the numbers of American deaths that might be caused by various levels of nuclear war. The lowest worth mentioning, for a "small" war, was two million.[9] By contrast, the number of deaths, American and other, from the attacks of 9/11 was less than 3,000, and the total number of deaths in terrorist attacks from 1970 to 2005 (excluding the terrorists

---

[9] Kahn [1961a] p.20. The bottom legend on the table was "Will the Survivors Envy the Dead?" a line that has become famous.

themselves) is estimated at less than 4,000.[10] That number excludes attacks in both Israel and Iraq, but no matter what might be added in, the sum will be a lot smaller than two million. Even if terrorists were to smuggle in a "dirty weapon" or even a small nuclear bomb or a major biological or chemical weapon, the potential number of deaths would remain far smaller than that.

A minimum of two million deaths, together with all the other economic and social destruction detailed by Kahn, was worth spending a lot to avoid, and was worth making major changes in policy. The dollars were spent and the policies were changed.

The numbers and times now are different. How much should the far smaller maximum numbers at risk today increase U.S. expenditures and America's way of life including Constitutional and other protections? Our own response so far has not been an entirely rational one: we have devoted substantial resources to fighting the insubstantial threat of non-nuclear terrorism. To a large extent we have deterred ourselves. We may even have been overcome by fear itself.

2. Theory

Our basic theory follows in the footsteps of Hirshleifer's work on both conflict and deterrence.[11] We study the problem of a friendly country deterring inimical activity by an enemy. We denote by $a \in [0, \bar{a}]$ the level of inimical activity conducted by the enemy. This consists of harmful activities ranging from terrorist attacks to full-scale invasion. The incentives of both sides are summarized by utility functions. The enemy receives a non-decreasing level of utility $u^e(a)$ from this activity; the friendly country receives a non-increasing level of utility $u^f(a)$. We may normalize friendly utility to be non-negative, so that $u^f(a) \geq 0$.

To model deterrence, we model the friendly country as having the capability of committing to a retaliation at level $r(a)$ in response to the inimical activity level a. If retaliation takes place, it represents a cost borne by the enemy, so that the overall enemy utility is $u^e(a) - r(a)$. Retaliation also has a cost to the friendly country: imposing r on

---

[10] The primary source of the estimate is U.S. State Department [2003] supplemented by public sources for the most recent years.

[11] See footnote 3.

the enemy costs cr, where $c > 0$ is the marginal cost of retaliation. Consequently, the overall friendly utility is $u^f(a) - cr(a)$.

Given perfect commitment and unlimited ability to retaliate, the optimal commitment plan is to commit to a sufficiently high level of response to any inimical activity that all inimical activity is deterred. The great advantage of this is that because inimical activity does not take place, the cost never has to be paid. This is a well-known basic result on deterrence, discussed extensively by Schelling [1960] and a simple elegant model can be found in Hirshleifer [1989].

This result however, assumes a degree of perfection – that there is a single monolithic enemy that does not make mistakes or miscalculate, and that the friendly country of a is assumed to observe a perfectly, that is, the response is certain. As a model of "bounded rationality" of the enemy and of noise in the signal of their activity level, we use the quantal response model of McKelvey and Palfrey [1993]. That is, we assume that the enemy is less likely to carry out activities that yield low utility than high utility. Nevertheless, because of random elements of preferences or miscalculation, there is a positive probability they will carry out inimical activity that would not be profitable given the base utility function. The quantal response model can also capture the possibility that there is not a single monolithic enemy, but a diverse array of enemies with different motives.

Specifically, we assume that the probability density that the enemy carries out an inimical level of activity a has the logistic form

$$f(a) = \frac{e^{\lambda(u^e(a) - r(a))}}{\int_0^{\bar{a}} e^{\lambda(u^e(a') - r(a'))} da'}.$$

Here λ is a parameter that measures the "rationality" or the inverse of the "noisiness" of the enemy. If λ is zero, the enemy behaves completely randomly – regardless of utility and cost, his play is described by a uniform distribution. If λ is very large, then he puts virtually all probability on a narrow range of alternatives near the utility maximizing alternative. This is a variation of the multinomial logit model of McFadden [1973] and has been widely and successfully used to analyze choices ranging for decisions over urban transportation (for example, Train [1980]) to the choice of a life of crime (for example, Goldberg and Nold [1980]).

The problem of the friendly country is then to choose the retaliation function to maximize the objective

$$U^f = \int_0^{\bar{a}} [u^f(a) - cr(a)] \frac{e^{\lambda(u^e(a)-r(a))}}{\int_0^{\bar{a}} e^{\lambda(u^e(a')-r(a'))} da'} da$$

If we differentiate this with respect to $r(a)$, we find that the sign of the derivative is given by

$$\frac{dU^f}{dr(a)} \doteq -[u^f(a) - cr(a)] - (c/\lambda) + U^f$$

where $\doteq$ means has the same sign as. Notice that the third term is an aggregate that does not depend on $r(a)$, while the first two terms are negative when $r(a) = 0$ and are increasing in $r(a)$. The only possible optima are $r(a) = 0, \bar{r}$, that is, either do not retaliate, or retaliate at the maximum possible level. Since $u^f(a)$ is decreasing in a, it follows that below some threshold a^* there should be no retaliation, and above that level, relation should be set at $\bar{r}$.

To understand this solution and its limitations, it is helpful to work through a simple example and consider why the simple threshold rule that is optimal in the formal model might or might not be optimal in a broader context. It may be helpful to consider the common context of ordinary criminal activity. Suppose that there are three levels of criminal activity – none, kidnapping or murder. Suppose that there is no penalty if there is no crime and that penalty for murder is the death penalty. How should the penalty for kidnapping be set? Conventional reference to the theory of marginal deterrence says that the penalty should be set below the death penalty because otherwise kidnappers would switch to murder, there being no difference in penalty between the two. But this analysis misses the fact that increasing the penalty for kidnapping from life in prison to the death penalty would not only encourage some kidnappers to commit murder, it would also encourage some kidnappers to stop committing crimes altogether.

For simplicity we might imagine that there are two types of criminal – kidnappers who either commit no crime or kidnap, but who never murder, and murderers who either kidnap or murder, but always commit a crime. Suppose that if there is no penalty for kidnapping, both kidnappers and murders choose to kidnap, while if there is a death

penalty for kidnapping, kidnappers choose not to commit crimes and murderers choose to murder. When would it make sense to choose an intermediate level of punishment – life in prison – rather than one of the two extreme punishments for kidnapping?

Consider what happens if we were to raise the penalty on kidnapping from no penalty to a modest penalty. Suppose first that kidnappers are relatively indifferent between no crime and kidnapping and that murderers are relatively indifferent between kidnapping and murder. Then a small penalty for kidnapping will tilt the balance for the nearly indifferent kidnappers, convincing them to stop committing crimes, while the relatively indifferent murderer prefers to go on kidnapping because the penalty is so much lower than it is for murder, and murder is not all that much better than kidnapping. So the modest penalty causes the kidnappers to stop kidnapping and the murders to stop murdering – the best outcome we can hope for. Here an intermediate penalty is socially desirable.

By way of contrast, suppose that kidnappers like kidnapping very much, and that murders like murder very much. Then a small penalty for kidnapping will not dissuade the kidnappers from kidnapping, since they enjoy it so much more than not committing any crime. However, the murderers, who much prefer murdering to kidnapping will be persuaded by a modest penalty for kidnapping to switch to their preferred crime. In this case an intermediate penalty results in the kidnappers kidnapping and the murderers murdering – the worst possible outcome that can occur.

Turning back to the arithmetic of deterrence, in our expression for the optimal value of $r(a)$ the net benefit of individuals who are deterred from choosing the activity level a is measured by U^f, the expected utility from their switching to other crimes. In the logit model, this utility is independent of $r(a)$ – that is, in the logit model, it is implicitly assumed that the alternatives that are switched to by discouraged kidnappers are independent of the punishment for kidnapping. This neutrality assumption seems a plausible starting point for analysis, especially as the logit model has proven a good fit to such a variety of data. However, the model also direct our attention to the circumstances under which partial rather than all or nothing deterrence would be desirable: if U^f, the utility from alternative activities is decreasing in $r(a)$ – that is, small levels of punishment convince kidnappers to switch but not murderers, then intermediate levels of punishment are favored. If the utility from alternative activities is increasing in $r(a)$ –

that is, small levels of punishment convince murders to switch but not kidnappers, then the all-or-nothing result is reinforced.

Turning back to international deterrence, in addition to the possibility that increased penalties change the mix of types who change their activity, there are several other reasons why all-or-nothing punishment may not be optimal.

1. As the enemy gradually ramps up his level of inimical activity, he observes that there is no response, so he may not learn that there is a threshold until it is too late.
2. Maximal retaliation in response to a minor offense may not be credible.
3. Maximal retaliation may create perverse incentive for third parties.

The first possibility becomes less important as either of the other two effects become significant. In addition, in a noisy world, enemies sometimes would cross the threshold, so they will learn where it is.

The issue of credibility is an important one; the cost of failing to punish a particular level of inimical activity, however, is not merely that the enemy will continue to engage in that activity in the future. Potentially the cost is that all reputation is lost, and the enemy will no longer believe that the friendly country will engage in retaliation for any level of activity. Indeed, it has been argued that U.S. response to the Lebanon barracks bombing in the Reagan administration and the killing of troops in Mogadishu during the Clinton administration had exactly the effect that it convinced our enemies that we would not retaliate in response to provocation, and so made inevitable the September 11 attack.

Credibility can also be endangered from the other side, however. Going ahead with a punishment because of fear that failure to do so would endanger our credibility may end up being so costly that a future enemy will believe that we will not do that again. This is one possible outcome of the unanticipated costs and difficulties to the United States of the attack on Iraq.

The large-scale loss of credibility as a matter of theory is prototypical of reputational models such as Milgrom and Roberts [1982], Kreps and Wilson [1982] and Fudenberg and Levine [1989]. The implication of this theory in our setting is that credibility lowers the relevant value of $\bar{r}$. Specifically, the loss from failing to carry out

the threat is $u^f(\overline{a}) - U^f$, where U^f is the solution to the commitment problem – and this is regardless of the level of activity a that must be responded to. So credibility has the implication that $c\overline{r} \leq \left[u^f(\overline{a}) - U^f \right]/i$ where i is the interest rate, but no more than that. Notice that in the case of the modern terrorist threat $u^f(\overline{a}) - U^f$ is far lower than it was during the cold war, so the level of retaliation we can credibly carry out is much lower.

Finally, there is the issue of incentives for third parties. Suppose that we committed to invading Iran if shots were fired from its embassy at spectators – as actually happened in the Libyan Embassy in London in 1984. Iran of course has many enemies besides us – Israel, the Iraqi Sunnis, the Kurds, Iranians in exile – to name a few. Naturally any of these enemies would be then be eager to stage an incident at the Iranian embassy involving shots at spectators – the "terrorist" attack on the Iranian embassy in London staged by Iraq in 1980 being a possible model for such an operation. Such "false flag" efforts are not uncommon in intelligence operations. Despite compelling evidence to the contrary, it is widely believed in the Arab world that the September 11 attacks in the U.S. were staged by Israel, and there is the possibility that some of the car bombings carried out in Lebanon were in fact staged by Israel. Although it did not involve third parties, World War II began with a false flag operation – German commandos dressed in Polish uniforms assaulted a German radio station on the German/Polish border. The false intelligence on Iraqi WMD provided by Ahmed Chalabi and the Iraqi National Congress is current example.

We face then a trade-off. If we use severe penalties, we deter inimical activity by our enemies – but encourage inimical activity by our enemies' enemies. By moderating penalties for lesser offenses, we reduce the incentive of third parties to manipulate our retaliation to their advantage. The effect of this in the noisy world is complicated to analyze, but analysis of the simple world of perfect decision making and information gives some hints. Suppose first that a maximum retaliation takes place after a threshold a^* is reached. If the first party can reach the threshold at sufficiently low cost that the benefit of the retaliation to the friendly party exceeds the cost, then the only possible equilibrium is for full retaliation to take place. The enemy, recognizing this, will engage in the maximum level of inimical activity $\overline{a}$ resulting in the worst possible result – no deterrence, and full cost of retaliation. If, however, retaliation is limited so that the cost to the enemy is simply equal to the benefit he derives from the inimical activity, a third

party may find that there is no level of activity for which the benefit to it exceeds its cost. This has the further implication that the enemy must be cautious to keep inimical activity low enough that the third party is not tempted to trigger increased retaliation. Here scaling the retaliation to the level of gain of the enemy discourages third party intervention, while the possibility of that intervention further reduces inimical activity by the actual enemy.

With these insights, let us return to the simple model and examine the comparative statics with respect to the parameters: λ, the level of rationality, $\overline{r}$, the level of retaliation and $\overline{a}$, the maximum level of the enemy attack. Our goal is to discover the implications of moving from a Cold War to a War on Terrorism for deterrence and utility.

We start with the level of utility U^f. Clearly increasing $\overline{r}$ can only help. By the usual type of envelope theorem result, it suffices to consider the partial derivative with respect to the parameters. Increasing $\overline{a}$ lowers friendly utility by adding less desirable outcomes. The case of λ is more complicated, but provided that $u^e(\overline{a}) - \overline{r} \leq u^e(0)$, so that deterrence is sufficient to prevent the worst possible outcome, increasing λ increases the weight on the favorable alternatives below $a*$ and reduces the weight on the unfavorable alternatives above $a*$, and so increases friendly utility.

Turning to the impact of the parameters on $a*$, the change will be the opposite direction of the change of the partial derivative of $-[u^f(a) - c\overline{r}] - c/\lambda + U^f$, holding fixed the optimal strategy. Changing $\overline{r}$ holding fixed the optimal strategy increases this partial derivative; changing $\overline{a}$ decreases this partial derivative through the effect on U^f, and increasing λ increases the partial derivative.

We summarize the comparative statics in the following table.

	λ	$\overline{r}$	$\overline{a}$
U^f	+	+	-
$a*$	-	-	+

Our next goal is to compare the cold war with terrorism
- ➢ We expect λ to decrease reflecting a less rational, more heterogeneous enemy
- ➢ Terrorists are vastly less capable of harming us, meaning a decrease in $\overline{a}$.

➢ The decreased possibility of harm greatly increases our utility from the most harmful enemy activity, and so should decrease our maximum credible retaliation $\bar{r} = \left[u^f(\bar{a}) - U^f \right] / ci$.

The net impact of moving from a Cold War environment to a war on terrorism on our welfare U^f and the optimal threshold for retaliation a^* is ambiguous. Lower rationality of the opponent λ and a lower retaliation $\bar{r}$ lower our welfare U^f and should lead us to be more tolerant of inimical activity as measured by a higher value of a^*. On the other hand, the greatly reduced possibility of harm $\bar{a}$ raises our welfare and should lead us to be less tolerant of inimical activity.

The intuition for λ and $\bar{a}$ run like this. On the one hand, our enemy is less rational and more random. That hurts us, and means that there is less reason to retaliate – that is we should tolerate more inimical activity. On the other hand, our enemy is vastly less capable of harming us. That helps us. It also means we should be less tolerant of inimical activity – we need not be so concerned that the threat of retaliation for small attacks will drive terrorists to much larger attacks because they are not capable of carrying out much larger attacks.

And finally, we should stress that the model can run in either direction. That is, the same formulations can be used to examine the deterrence of a friendly country by an enemy.

3. Conclusions

Our four major conclusions are that:

➢ During the cold war, mutual deterrence between the United States and the Soviet Union was very strong.
➢ Deterrence of terrorist states is similar to cold-war deterrence in many respects.
➢ Deterrence of terrorist groups, however, is different and more difficult.
➢ The costs of a failure to deter terrorists will be far lower than they would have been during the cold war.

The Cold War

The cold war provided almost a textbook example of the operation of the deterrence model. For two levels of potential Soviet activity (*a*) inimical to the United States—a nuclear attack on the U.S. homeland and an attack, nuclear or not, on our West European allies, U.S response was almost certain, and expected to be as close to devastating as possible or necessary: $\bar{r}$ was very high. Below that, for lesser communist actions—subversion or aggression—the likelihood of military retaliation against the Soviet Union was close to zero. The dichotomy is consistent with the all-or-nothing implications of the concept.

A nuclear attack on the U.S. (the only kind that could have made any military or logistical sense) would have been met by an overwhelming nuclear response; the strategic doctrines instituted by Defense Secretary McNamara created guaranteed the preservation of sufficient American forces to mount such a response. A non-nuclear attack on Europe might be met initially by a non-nuclear defense, but if that defense were swept aside—as seemed likely—escalation, perhaps starting with "small" tactical nuclear weapons, seemed probable; and that breach in the firebreak would open the road to all-out nuclear exchange. The certainty of the response to an attack in Europe was less than the certainty of a response to an attack on the U.S., but it was high enough to deter such Soviet aggression (although France worried enough about American "resolve" that it created its own independent nuclear force.) In any case, neither of these Soviet activities ever took place: $\bar{r}$ was high enough, given the high utility u of Russian society and leadership with a lot they wanted to conserve, and the coherence λ of a centralized system, to deter them completely.

The same was not true, however, for Soviet communist-sponsored subversion, or for the direct Soviet aggression in Afghanistan toward the end of the cold war. For these, U.S. r was close to zero, and thus, as suggested by the model, these activities were carried out with impunity.

The model also worked in reverse. The U.S. was deterred from military attack on the USSR or its Warsaw Pact allies, but not from lesser activities such as the support of the "Contra" attack on the pro-Soviet Nicaraguan regime. The result of the two-directional applicability of the model was the stability of "Mutual Assured Destruction".

Terrorist States

In the war on terrorism, although the United States accuses some states, notably Syria and Iran, of harboring or even sponsoring terrorist groups, that is not the issue here. The real fear is of states obtaining nuclear weapons or perhaps other WMD, and threatening to use them and perhaps actually using them for political ends. The two current candidates are North Korea which may well have a few nuclear weapons, and Iran which is widely believed to be on its way to getting them.

Deterring the use of those weapons seems possible, although not easy. The coherence parameter λ is about as high as it can get for the apparent one-man dictatorship of North Korea; Iran is more complex—a somewhat chaotic somewhat populist nation governed mostly by a single "supreme authority"—but with nuclear programs seemingly under tight control. National utility is probably high for Iran, with substantial resources, some wealth, and a structured society; North Korean society has little to do with national decision-making, which is entirely in the hands of the "dear leader" but his desire to retain power may also put his potential losses at a high level.

High λ and the potential of a large loss, together with the credibility of an overwhelming U.S. response to the use of nuclear weapons and the military capability to implement such a response with or without its own nuclear weapons, mean that deterrence is strong. That is true even though the stakes for us are much lower than they were in the cold war. What counts for deterrence is the stake for the enemy to be deterred, and for a terrorist state, that will be continued existence.

One of the authors of this paper has proposed a system of deterrence by the major nuclear powers to prevent nuclear first use.[12] That may not be possible politically, but the United States has the unilateral power to implement such a system.

Terrorist Groups

For four reasons, deterrence of non-state terrorist groups like al-Qaeda will be more difficult.
1. The extremely diffuse nature of terrorist decision-making. Al Qaeda itself is a loose network in which bin Laden and his associates have influence but no command over the cells, and al-Qaeda is merely the most notorious organization

---

[12] Levine [1993].

in a loose network of networks. The result is a very low λ, in which enemy behavior as seen from our point of view is near random.

2. Terrorist cells are mobile and elusive, which makes targeting and retaliation, and threats to retaliate, less than inexorable.

3. The value systems of some of those to be deterred are at best difficult to discern, at worst difficult to threaten. Ironically, those terrorists who are basically thugs should be able to estimate their potential gains and losses in ways that makes deterrence easy. But those who are ideologues—in the limit, suicide bombers— seem to welcome what we might consider punishment. This may even be objectively true if retaliation encourages revenge: that, it is apparent that punishment of those we consider terrorists in Iraq has enraged other Muslims, inspiring them to become terrorists.

4. It is not credible to threaten huge retaliation against relatively the small actions of which a terrorist group may be capable. It is not relevant, at least, for a liberal democracy, although Hitler and Saddam did it.

On the other hand, the damage that can be done by these groups without state backing is relatively small. Even another September 11 level attack would scarcely be a threat to Western civilization. The real threat is that a state – Iran, Pakistan or North Korea – might by design, or even by accident, provide a terrorist group with a nuclear weapon or weapons. That threat is subject to deterrence, since the deterrence can be aimed against the state. Indeed, we are in the process quite correctly of publicizing our ability to trace nuclear residue to a particular nuclear program. The clear implication is that a nation that we believe guilty of providing a nuclear weapon to a group that uses it against us can expect fearsome retaliation. Notice that here our deterrence provides incentives not only against intentional provision of a nuclear weapon, but also accidental provision, for example through lax security. The unfortunate weak link in the deck is Russia. While we might credibly retaliate against Iran, Pakistan or North Korea in response to a "small" nuclear attack involving one or a few weapons, it is less credible that we would do so against Russia which still maintains a nuclear arsenal capable of far greater than a "small" nuclear attack.

The Costs of Failure

As noted, the model works in both directions—a friendly country deterring an enemy and an enemy deterring a friend. That resulted in the symmetrical Mutual Assured Deterrence of the Cold War, based on fear of numbers like Kahn's minimum two million deaths from a "small" nuclear exchange. The "War on Terror" is not symmetrical, but for the United States it is, or should be, far less fearful.

Looking first at terrorist states, it is conceivable that sometime in the not-too-distant future North Korea, or less likely given the geography Iran, could achieve the capability of launching a nuclear missile at the United States. Since the U.S. could and quite possibly would retaliate with an all out punishment, wiping out the entire offending state, deterrence, even against a megalomaniac concerned only with his own power, is likely to work. Even if deterrence were to fail, however, and even if the enemy were lucky enough to hit a population center (the Bush administration's anti-ballistic-missile program is intended to decrease that probability; whether the potential decrease is worth the cost is outside the scope of this paper), the small size of a nuclear weapon deliverable by a terrorist state would be far below the lower end of the cold-war scale, nor could it escalate to a thermonuclear holocaust.

Successful action by a non-state terrorist group is more likely but is likely to be much smaller. But the potential number of casualties, even from a Weapon of Mass Destruction is nearer that from the 9/11 attacks than it is to a cold-war scale. The Office of Technology Assessment of the U.S. Congress, for example, estimates that "Effectively disseminated, a single release of a chemical or biological weapons could cause tens of thousands of casualties."[13] Estimates for radiation weapons are lower.[14] The most fearsome possibility would be an actual nuclear explosion. The Center for Defense Information calculates that "A small bomb, say 15-kilotons, detonated in Manhattan could immediately kill upward of 100,000 inhabitants, followed by a comparable number of deaths in the lingering aftermath."[15]

---

[13] U.S. Congress, Office of Technology Assessment [1993].

[14] The Federation of American Scientists, for example, estimates that for the kind of weapon that might be assemblable by terrorists, "near-term casualties would likely be few and the increased cancer risks in the exposed population slight" von Huppel [2001].

[15] Center for Defense Information Terrorism Project [2001].

These are terrible numbers but the largest is still an order of magnitude smaller than Kahn's smallest. The cold-war threat of deaths from the tens of millions to the hundreds of millions brought about immense expenditures on strategic and other weapons. The changes in America's way of life, exemplified by the Kennedy administration's proposals for extensive civil defense efforts, soon petered out; the threat simply did not seem great enough. The excesses of the McCarthy era were directed at internal political subversion, not the external Soviet military threat; terrorism was not a word used or thought of.

In contrast, the smaller dangers of 21st century terrorism may threaten major inroads on American liberties. The attacks of 9/11 shifted the always-unstable balance between "security" and "civil liberties" toward the security end; public opinion is divided almost equally between the two, considered in the abstract. But the next terrorist attack, whether at the 9/11 scale or ten times as large but still 100 times smaller than the "small" cold-war attack, may move the public and its representatives decisively in the direction of security at all costs, including the cost of liberty.

Looking across the range of potential terrorist threats, by nations or by groups, deterrence of all hostile actions is more complex and less certain than was cold-war deterrence of nuclear and other major attacks. Nor is prevention certain; it seems likely that one attempt or another will succeed, and it may well result in more American casualties than 9/11. But it seems certain that the casualties will be far fewer than the smallest nuclear exchanges postulated for the Cold War.

The real danger to the United States is fear itself, resulting in major changes to our "way of life", both in anticipation of such an attack, and in response to it.

References

Center for Defense Information Terrorism Project [2001], *What if the Terrorists Go Nuclear?* Updated October 1.

Goldberg, I. and F. C. Nold [1980], "Does Reporting Deter Burglars?--An Empirical Analysis of Risk and Return in Crime," *The Review of Economics and Statistics*.

Fudenberg, Drew and David K. Levine [1989], "Reputation and equilibrium selection in games with a patient player," *Econometrica*, **57**: 759-778.

Hirshleifer, Jack [1949], *Tasks in Measuring Economic War Potential,* RAND RM-0301/PR.

Hirshleifer, Jack [1963], *Disaster and Recovery: A Historical Survey,* RAND RM-3079.

Hirshleifer, Jack [1989], *The Dimensions of Power as Illustrated in a Steady-State Model of Conflict* (RAND N-2889).

Kahn, Herman [1961a], *On Thermonuclear War,* Princeton University Press.

Kahn, Herman [1961b], "The Delicate Balance of Terror", *Foreign Affairs,* April.

Kahn, Herman [1962], *Thinking About the Unthinkable,* Horizon Press.

Kreps, D. and R. Wilson [1982], "Reputation and Imperfect Information," *Journal of Economic Theory,* **50**: 253-79.

Levine, Robert A. [1993], *Uniform Deterrence of Nuclear First Use,* RAND MR-231-CC.

McFadden, Daniel [1973], "Conditional logit analysis of qualitative choice behavior," Berkeley, Calif.: Institute of Urban and Regional Development.

McKelvey, R. and T. Palfrey [1995], "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, **10**: 6-38.

Milgrom, P. and J. Roberts [1982]: "Predation, reputation , and entry deterrence," *Journal of Economic Theory*, **27**: 280-312.

Schelling, Thomas C. [1960], *The Strategy of Conflict*, Harvard University Press

Schelling, Thomas C. and Morton H. Halperin [1961], *Strategy and Arms Control*, Twentieth Century Fund.

Train, K. [1980], "A Structured Logit Model of Auto Ownership and Mode Choice," *The Review of Economic Studies*.

U.S. Congress, Office of Technology Assessment [1993], *Proliferation of Weapons of Mass Destruction: Assessing the Risks,* OTA-ISC-559, August.

U.S. State Department [2003], *Significant Terrorist Incidents, 1961-2003: A Brief Chronology.*

von Huppel, Franz [2001], *Recommendations for Preventing Nuclear Terrorism,* FAS Public Interest Report, Volume 54, Number 6.